

SEQUENCE LISTING - WEE virus strain 71V-1658 - SEQ ID NO 1

LOCUS AF214040 11484 bp RNA VRL 27-JAN-2000
DEFINITION Western equine encephalomyelitis virus strain 71V-1658 nonstructural polyprotein and structural polyprotein genes, complete cds.
ACCESSION AF214040 AF143811
VERSION AF214040.1 GI:6760410
KEYWORDS
SOURCE Western equine encephalomyelitis virus.
ORGANISM Western equine encephalomyelitis virus Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.
REFERENCE 1 (bases 1 to 11484)
AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A., Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
TITLE Complete genomic RNA sequence of western equine encephalitis virus and expression of the structural genes
JOURNAL J. Gen. Virol. 81 Pt 1, 151-159 (2000)
MEDLINE 20109043
REFERENCE 2 (bases 1 to 11484)
AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A., Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1999) Medical Countermeasures Section, Defence Research Establishment Suffield, P.O. Box 4000, Stn Main, Medicine Hat, Alberta T1A 8K6, Canada
FEATURES Location/Qualifiers
source 1..11484
/organism="Western equine encephalomyelitis virus"
/strain="71V-1658"
/db_xref="taxon:11039"
5'UTR <1..24
misc_feature 1..11484
/note="sequence of this entry added to the sequence of AF214039, strain CBA87, constitutes the Western equine encephalomyelitis virus complete genome"

CDS 25..7428
/note="encodes nsP1, nsP2, nsP3 and nsP4 proteins"
/codon_start=1
/transl_except=(pos:5581..5583,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="AAF28339.1"
/db_xref="GI:6760411"

/translation="MERIHVDLDADSPYVKSLQRTFPQFEIARQVTNDHANARAFS
HVATKLIESEVDRDQVILDIGSAPVRHAHSNHRYHCICPMISAEDPDRLQRYAERLKK

SDITDKNIASKAADLLEVMSTPDAETPSLCMHTDATCRYFGS VAVYQDVYAVHAPTSI
YHQALKGVRTIYWIGFDTPFMYKNMAGSYPTYNTNWADERVLEARNIGLGNSDLQES
RLGKLSILRKRLQPTNKIIIFSVGSTIYTEDRSLLRSWHLPNVFHLKGKSNFTGRCGT
IVSCEGYVIKKITISPGLYGVENLASTMHREGFLSCKVTDTLGERVSFAVCTYVPA
TLCQMTGILATDVSVDDAQKLLVGLNQRIVVNGRTQRNTNTMQNYLLPVVAQAFSRW
AREHRADLDDEKELGVRERTLTMGCCWAFKTQKITSIYKKPGTQTIKKVPAVFDSFVI
PRLTSHGLDMGFRRRLKLLLEPTVKPAPAITMADVEHLRGLQQEAEVAAAEEIREAL
PPLLPEIEKETVEAEVDLIMQEAGAGSVETPRGHIRVTSPGEEKIGSYAILSPQAVL
NSEKLACIHPLAEQVLVMTHKGAGRKYVEPYHGKVIVPEGTAVPVQDFQALSESATI
VFNEREFVNRYLHHIAINGGALNTDEEYYKTVKTQDTDSEYVFDIDARKCVKREDAGP
LCLTGDLVDPFHEFAYESLKTRPAAPHKVPTIGVYGVPGSGKSGIIKSAVTKKDLVV
SAKKENCAEIIRDVRRMRRMDVAARTVDSVLLNGVKHPVNTLYIDEAFACHAGTLLAL
IAIVKPKVVLCGDPKQCGFFNMMCLKVHFNHDICTEVYHKSISRRCQTVAIVSTL
FYDKRMKTVNPCADKIIIDTTGTTKPHKDDLILTCFRGWVKQLQIDYKNHEIMTAAAS
QGLTRKGVYAVRYKVNEPNPLYSQTSEHVNLLTRTEKRIWKTLAGDPWIKTLTAKYP
GDFTASLDDWQREHDAIMARVLDKPQTADVFQNKVNVCWAKALEPVLATANI VLTRQQ
WETLHPFKHDRAYSPEMALNFFCTRFFGVLDSDLFSAPTVALLYRDQHWDNSPGKNM
YGLNREVAKELSRRYPCITKAVDTGRVADIRNNTIKDYSPTINVVPLNRRLPHSLIVD
HKGQGTTDHSGFLSKMKGKSVLVI GDPISIPGKKVESMGPLPTNTIRCDLDLGIPSHV
GKYDIIFVNVRTPYRNHHYQQCEDHAIHHSMLTCKAVHHLNTGGTCVAIGYGLADRAT
ENIITAVARSFRFTRVCQPKNTAENTEVLVFFGKDNGNHTHDQDRLGVVLDNIYQGS
TRYEAGRAPHYAVRVIRGDISKSADQAIVNAA NSKGQPGSGVCGALYRKWPAAFDRQPIA
VGTARLVKHEPLIIHAVGPNSKMPPEPEGDLKLAAYMSIASIVNAERITKISVPLLS
TGIYSGGKDRVMQSLHHLFTA FDTTDADVTIYCLDKQWETRIIEAIHRKESVEILDDD
KPV DIDLVRVHPNSSLAGRPGY SVNEGKLYSYLEGTRFHQTAKDIAEIHAMWPNKSEA

NEQICLYILGESMSSIRSCKPVEESEASAPPHTLPCLCNYAMTAERVYRLRSAKKEQF
AVCSSFLLPKYRITGVQKLQCSKPVLFGVVPPAVHPRKYAEIILETPPPPATTVIC
EPTVPERIPSPVISRAPSAESLLSLGGVSFSSSATRSSTAWSDYDRRFVVTADHQAN
TSTWSIPSAPGLDVQLPSDVTDSHWSIPSASGFEVRTPSVQDLTAECAKPRGLAEIMQ
DFNTAPFQFLSDYRPVPAPRRRPIPSRSTASAPPVPKPRRTKYQQPPGVARAISEAE
LDEYIRQHSNXRYEAGAYIFSSSETGQGHLQQKSVRQCKLQEPILDRAVHEKYYAPRLD
LEREKMLQKKLQLCASEGNRSRYQSRKVENMKAITAERLISGLGTYLSSEVNPVECYR
VNYPVPIYSSTVINRFTSAEVAVKTCNLVIQENYPTVASYCITDEYDAYLDMVDGASC
CLDTATFCPAKLRSYPKKHSYLPQEIERSAVPSPIQNTLQNVLAATKRNCNVTQMREL
PVLDAAFNVDCKKYACNDEYWDTFRDNPIRLTENVTVQYVTKLKGPKAAALFANTH
NLKPLQEIPMDQFVMDLKRDVKVTPGTKHTTEERPKVQVIQAADPLATAYLCGIHRELV
RRLNAVLLPNIHTLFDMAEDFDAIIAEHFHHGDPVLETDIASFDKSEDDAIAISALM
ILEDLGVDQPLLDLIEAAFGNITSVHLPTGTRFKFGAMMKSGMFLTFVNNTLVNIMIA
SRVLRERLTTSCAACASIGDDNIVHGVVSDTLMAERCATWLNMEVKIIDAVIGIKAPYF
CGGFILVDQITGTACRVADPLKRLFKLGKPLPVDDTQDCDRRALHDEAMRWNRIGHT
DELVKAVESRYEIILAGLIITSLSTLAESVKNFKSIRGSPITLYG"

CDS 7473..11183
/note="encodes nucleocapsid, E3, E2, 6K and E1 proteins"
/codon_start=1
/product="structural polyprotein"
/protein_id="AAF28340.1"
/db_xref="GI:6760412"

/translation="MFPYPQLNFPPVYPTNPMAYRDPNPPRRWRPFRPLAAQIEDL
RRSIVNLTFKQRSPNPPPGPPPQKKKSAPKPKPTQPKKKQQAKRTKRKPQPGKRQRM
CMKLESDKTFPIMLNGQVNGYACVVGGLMKPLHVEGKIDNEQLAAVKLKKASMYDLE
YGDVPQNMKSDTLQYTSKPPGFYNWHGAVQYENGRFTVPRGVGGKGDSGRPILDNR
GRVVAIVLGGANEGRRTALSVVTWNQKGVTIRDTPEGSEPWSLVTALCVLSNVTFC

KPPVCYSLTPERTLDVLEENVDPNYDTLLENVLKPSRRPKRSITDDFTLTSPYLG
CPYCRHSTPCFSPIKIENVWDESDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDHDH
DIKEDSMEKIAISTSGPCRRLGHKGYFLLAQCPGDSVTVSITSGASENSCTVEKKIR
RKFVGREEYLFFFVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
KPPSGKNVTVYECKCGDYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSPDLIRHT
DHSVQGKLHIPFRLPTVCPVPLAHTPTVTKWFKGITLHLTAMRPTLLTRKLGLRAD
ATAEWITGSTSRNFSVGREGLEYVWGNHEPVRVWAQESAPGDPHGPHEIIIIHYYHRH
PVYTVIVLCVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT
NAETFGETLNHLWFNNQPFLWAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAF
EHATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPSTNKEYVTCKFHTVIPSPQ
VKCCGSLECKASSKADYTCRVFGGVYPFMWGGACQFCDSENTQLSEAYVEFAPDCTID
HAVALKVHTAALKVGLRIVYGNTTAHLDTFVNGVTPGSSRDLKVIAGPISAAFPFDH
KVVIRKGLVNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHV
TOAVSGYEMWKNNSGRPLQETAPFGCKIEVEPLRASNCAYGHIPISIDIPDAAFVRSS
ESPTILEVSVCTVADCIYSADFGGSLTLQYKADREGHCPVSHSTTAVLKEATTHVTAV
GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHIIGEPHKVDQEFQAAVSKTSW
NWLLALFGGASSLIVVGLIVLVCSSMLINTRR"

3' UTR

11184..11484

BASE COUNT 3251 a 2842 c 2808 g 2583 t
ORIGIN

1 accctacaaa ctaatcgatc caatatggaa agaattcacg ttgacttaga tgctgacagc
61 ccgtatgtca agtcgttaca gcggacgttt ccacaatttg agatcgaagc aaggcaggc
121 actgacaatg accatgcca tgccagagcg ttttcgcattg tgcaacaaa gctcatttag
181 agcgaagtgc accgggacca agttatcttgc gacattggaa gtgcgcggcgt cagacatgca
241 cattccaatc accgttatca ttgtatctgc cctatgataa gcgtgaaga cccggacaga
301 ctacaacggt atgcagaaag acttaagaaa agtgcacatta ccgacaagaa catagcctct
361 aaggcggcag acctgcttgc agtcgttaca acaccagacg cagagactcc atctctgtgt
421 atgcacacag acgcacacgtg taggtacttt ggaagtgttag cagtatacca agatgtgtac
481 gcagtccatg caccgacatc aatctaccac caggcgctta aaggagttag gacaatttac
541 tggataggct ttgacacgac ccctttatg tacaataaca tggcaggttc ctaccctact

601 tacaacacga actgggctga cgagagagta ttggaagcac gtaacattgg cctcggtAAC
661 tcagatcttc aggagagcag gcttggaaaa ctctcaatcc ttaggaagaa gaggctccaa
721 cctactaata agatcatatt ctcgggtgg tcaacaatct acacagaaga tagatcaactg
781 ttacgttagct ggcacatctcc aaacgtgttc cacttggaaag gaaagtctaa cttcacaggt
841 agatgtggga ccattgtcag ctgtgaaggg tacgtcatca aaaagataac gatcagccc
901 ggactatacg gtAAAGTTGA gaacttggcg tccacaatgc atcgcgaggg ttttttgagt
961 tgccaaagtca cagatacgct gcgcggcgag agggttctt ttgctgtgt tacgtatgt
1021 ccagccacac tttgcgatca gatgacaggg attctggcaa ctgacgttag tggatgtac
1081 gcacaaaaac tattgggtgg gctcaaccaa aggattgtcg tcaatggtag gacgcaaaga
1141 aatactaaca caatgcagaa ctatctatta ccagtgtcg cccaggcggtt ttccagggtgg
1201 gcgcgtgaac atcgtccga cttggacgac gagaaagaac taggggtgcg ggagcgcact
1261 cttactatgg gctgctgctg ggcttcaag acccagaaaa tcacatccat ctacaagaag
1321 cttggtagc aaacaattaa gaaagtacct gccgtcttg actcatttgc gattccacgc
1381 cttaccagcc acgggctcga tatggcttc cggcgtaggc tcaagctgt gctgaacca
1441 actgtcaaac ccgcaccggc tattacaatg gcgcgtgtgg agcatctgcg tggcttacag
1501 caagaagctg aagaagtggc tgcagcgaa gagatcagag aagccctgcc acccttgc
1561 cctgaaatag aaaaagagac cgtagaggca gaagttagacc tcattatgca agaggcagga
1621 gcaggttagcg tggagacacc acgaggacac atcagggtga caagttaccc aggcgaagag
1681 aagattgggt cttacgctat actttcaccc caggcggtat tgaatagtga aaaactggcg
1741 tgtatccacc cattggcgga acaagactg gtaatgactc acaaaggtag ggcagggaga
1801 tacaaggtagc agccatacca cggtaaggc attgtaccag aaggacggc ggtccctgtt
1861 caagacttcc aggacatttag tgagacgc acgatcggtt tcaacgagag ggagttcgta
1921 aacagatacc tgcaccacat cgcaatcaac ggaggacgc taaacactga cgaagagtac
1981 tataagactg taaagactca ggacacagac tcagaatacg tcttcgatat tgacgcacga
2041 aagtgtgtta agcgagaaga cgcagtc cttgtgcctaa cccgtgatct ggtagatcca
2101 ccatttcacg agtttgcgtc cgagagtctc aagacacgac cagcagcacc tcacaaagtc
2161 ccaaccatcg gagtctatgg agtgcaggt tcaggtaaat ctggaatcat caaaagcgct
2221 gtgactaaga aagatctggt tgtgagtgcg aagaaggaaa actgcgcaga aatcatcagg
2281 gatgtaagga ggatgagacg tatggatgtt gctgcttaga ctgtcgattc agtgcattcta
2341 aatgggtta agcaccccgta taacactctg tacattgtg aggcatttgc ctggccatgca
2401 gggacgctgc tggcactgat tgccatcgta aaacctaaga aagtggattt gtgcggggac
2461 cccaaaacaat gcggcttctt taacatgatg tgcctgaaag tacattttaa ccatgacata
2521 tgcactgaag tgtaccataa aagcatctct aggaggtgc cacagactgt aaccggccatc
2581 gtctccacgc tcttctacga caagcgaatg aagacggta acccatgtgc tgataaaaatc
2641 atcatagata ccacaggac cacaaggccg cacaaggatg atctgattt aacctgtttc
2701 agaggatggg taaaacagct acagattgac tacaatccat acgaaatcat gactgcggct
2761 gcatcgcaag gacttacgcg gaaaggcggt tatgctgtca ggtacaaagt caacgagaat
2821 ccactctact cgcagacttc tgagcacgtg aacgtgttac ttacacgc acggaaacgc
2881 attgtctggg agacgcttagc tggatgtccc tggataaaga cacttacagc taaatatccc
2941 ggggatttca cggcttcatt ggacgactgg cagcgcgaac acgacgccc tatggcacgc
3001 gttcttgata agccgcagac agctgatgtg ttccagaata aggtgaacgt ctgtggcg
3061 aaggctttag agccagtctt ggccacggcc aacattgtgc tgacgagaca gcagtggag
3121 acgttgcacc cattcaagca tgacagagcg tactcacctg aaatggact gaacttctt
3181 tgcaccaggc tcttggagt agacctggac agtgggttat ttccgcctcc taccgtcgca
3241 cttacttaca gggatcagca ctgggataac tcgcccaggaa agaacatgt tggcttaat
3301 agagaggtag caaaggagtt gtcacggcga tatccgtgca tcacaaaagc ggttgcacaca
3361 ggcagggtag ctgtatataag gaataatacc atcaaggact actctccaac aattaatgtg
3421 gttccattaa atcgcgggtt gccccactcg ttgatgttgc accacaaaagg acagggtaca
3481 actgatcaca gcggattcct atctaagatg aaggcataat ctgtgttgc gatcggcgat
3541 cctatcagca ttccaggaa gaaagtagag tccatgggtc cattgcccac taataccatc

3601 aggtgtgatc tcgatttggg aatacctagc catgtcggt aatatgacat tatctttgtc
3661 aatgttagga ccccgtacag gaaccatcac taccaacagt gcgaggatca cgctatccac
3721 cacagcatgc taacgtgtaa ggctgtccac cacctgaaca ctggcggAAC atgtgtggct
3781 atagggtatg ggcttgctga tcgcgcAAC gagaatatca tcactgcggg ggcacgctca
3841 ttttaggttta cccgtgtctg tcagcctaag aacactgccc AAAAatactga ggttcttcc
3901 gtgttcttcg gcaaggacaa cggcaaccac acacatgacc aggacagact cggtgttagt
3961 ctgtacaaca tctatcaagg gtcaaccagg tacgaggcag ggagagctcc agcgtacaga
4021 gtgatcagag gtgacattag caagagcgct gaccaagcta tcgttaatgc tgctaatagc
4081 aaaggtcaac cagggtccgg agtgtcggt gcactgtacc gaaaatggcc ggctgcttt
4141 gatagacagc caatagctgt cgggacggct agacttgta agcacaacc gctcatcata
4201 catgctgtag gacccaaattt ttcttaagatg ccggAACCG agggcgcac taaagctcgca
4261 gctgcctaca tgagcatagc gtccatcgTC aacgctgagc ggattacaaa aatatcagta
4321 ccgtactgt caaccggcat ctattctggt ggcaaaagatc gagtgtatgca atcattgcat
4381 cacctgttca ctgcttcga cactacggat gcccgtatgtca ccatatattt ctggataaaa
4441 caatgggaga ccaggataat cgaggccatt caccgcaaag aaagcgtcga aattctggat
4501 gatgacaagc cagtagacat tgacttggc agggtccacc caaacagctc tttggcaggc
4561 agaccagggt actccgtcaa tgagggcaag ttgtattcat acctggaaagg tacacgattc
4621 catcagaccc ccaaggacat tgccgaaatc catgcaatgt ggcccaacaa atctgaggct
4681 aatgagcaga tttgttgcA catcctgggg gagagtatgt ccagcatccg ctccaaatgc
4741 ccagtagagg agtcagaggc gtctgtcca cctcacacac ttccatgcct gtgttaattac
4801 gctatgacgg ctgagcgcgt atacaggttgcg cgtctgcga agaaaagaaca gttcgccgt
4861 tgctcatcat tcctgttgcA gaagtacagg atcacaggcg tgcagaagct acagtgcagc
4921 aaaccagtcc tgtttcagg cgtcgatcca ccggctgtac accccaggaa gtacgcccggaa
4981 ataattctag aaacgccacc accgcccggca acgacaaccg taatatgtga accaactgtg
5041 ccagaacgtA taccacgtcc ggtgatttct agagcaccaaa gtgcggaaatc actgctatcg
5101 ctggcggcg tctcggttctc tagctctgca acacgctcgt caaccgcctg gagcgactat
5161 gacaggcggt ttgtgttac agctgtatgtc catcaagcga acacgtctac gtggagcatc
5221 cctagtgcct ctgggttggA cgtccagctg cttctgcacg tcaactgtatc ccactggagt
5281 attccaagtg catcaggctt tgaagtgaga acaccatctg tacaggacat aactgcggag
5341 tgcgtcgaaatc ctcgtggact ggccgaaata atgcaagact tcaatactgc tcctttccag
5401 tttctttcgg actacagacc agtaccggca ccacggagac gccccatccc atcacctaga
5461 tcgacggctt ccgcacactcc agttccaaag ccacgcggaa ctaagtacca acaaccacca
5521 ggagtgcgtA gagcgtatctc agaagcggag ttggacgagt acatccgtca acactccaaac
5581 tgacggatgt aagcgggagc gtatatttc tcatcgaaagg caggccaaagg tcaccttcaa
5641 cagaaatcag tacgtcaatgt taaactacaa gaacctataat tggatcgggc cgtccatgag
5701 aagtattacg ccccgccct cgtatctcgaa agagagaaaa ttgtacagaa gaaactgc当地
5761 ttatgcgcct ctgaaggaaa tagaagcagg tatcaatcac gaaaagttaga aaatatgaaa
5821 gcaattacag cggagcgtact catttctggA ttgggcacat atctatcatc agaagtgaat
5881 cctgtcgagt gttacagagt caattatctt gtaccaatct actcgtcaac ggttaattac
5941 aggtttacat ctgcagaggt cgcggtaaa acgtgcacat tagttatcca agagaattac
6001 cctacagtag ccagttattg tataacagat gaatacgtatc cgtatcttgc catggggac
6061 ggcgcacgt gctgtctaga tacagccact ttttgcggc ctaaaactgtgag aagctaccca
6121 aagaagcata gctatttgcA gccagagata agatcaggccg tcccatcgcc tatacagaat
6181 acattacaaa atgtatttggc tgcagctact aaaaggaatt gcaacgttac ccaaattgc当地
6241 gaattacctg tcttagattc ggcggcattt aatgttgatt gttcaagaa atacgc当地
6301 aatgtatgtact gggatatac ctttcgtat aaccctattt ggcttaactac agagaacgtt
6361 acgcaatatg tgacaaaagct gaaaggccg aaagcagcag cattgtttgc gaatactcat
6421 aatctaaaac cgttgcagga gataccatg gatcaattcg tcatggatct aaagagagat
6481 gtcaaaggta ctccggcac gaaacataca gaggagccgc ctaagggtgca ggttattcag
6541 gctgcagatc cccttgctac cgcttacattt tgcggatcc atcgggaaatt agtccgtaga

6601 ctgaatgcgg tgcttctgcc aaatatccat actctttcg acatgtcagc ggaagatttt
6661 gatgcgatta ttgctgaaca tttccaccac ggcgaccagg tattggaaac ggacatcgcg
6721 tcgtttgata aaagcgaaga cgacgctatc gccatttcgg cgttgatgat cttgaggac
6781 ttaggtgtcg accaaccgct cttagattt atagaggcgg cgttcggcaa tatacacatct
6841 gtgcacctac ctacaggaac gaggtttaaa tttggtgcca ttagtgcata cggtagttc
6901 ttaacgctgt ttgtcaacac actagtcaat atcatgattt atagcagagt actacgtgaa
6961 cggtaacca cgtcagcgt cgccgcctct atcggcgacg ataacatagt gcatgggtc
7021 gtctccgaca ctttgatggc ggagagatgc gccacttggc tgaacatgga agtaaaaatt
7081 attgatgcag ttattggat caaagcaccc tacttctgtg ggggatttat cttggggac
7141 cagataacag gcacagcctg cagagtcgca gaccctctaa aaaggcttt taagcttgg
7201 aaaccattgc cagtcgatga tacccaagac tgcgaccgc gcccggact gcatgatgaa
7261 gcaatgcgt ggaacagaat tggattacg gacgagttag tgaaggccgt agaatccaga
7321 tacgagatca tactggcagg cctgatcatc acgtctctgt ccacgttagc cgaaacgctt
7381 aagaacttca agagcataag agggagccca atcaccctct acggctgacc taaataggtg
7441 acgttagata cacgcaccta cccaccggca gaatgtttcc ataccctcag ctgaacttgc
7501 caccagttt ccctacaat ccgatggctt accgagatcc aaaccctcct aggcccgct
7561 ggaggccgtt tcggcccccg ctggctgctc aaatcgaaga tcttaggagg tcgatagtca
7621 acttgacttt caaacaacga tcacctaatt cggcccgagg tccaccgc aagaagaaga
7681 agagtgcctc taagccaaaa cctactcagc ctaaaaaagaa gaagcagcaa gccaagagga
7741 cgaaacgcaa gcctaaacca gggaaacgac aacgtatgtg tatgaagttt gagtcggaca
7801 agacatttcc gatcatgctg aacggccaa tgaatggata tgcctgcgtt gtcggaggaa
7861 ggctgatgaa accactccac gttgaaggaa aaattgataa tgagcaatta gcggccgtga
7921 aattgaagaa ggctagcatg tacgacttgg agtacggcga cgttccccag aacatgaaat
7981 cagacacgct gcagtagcacc agcgacaaac caccggctt ctacaactgg caccacggcg
8041 cagtccagta tgagaatggg agatttaccg taccggagagg agtggccggg aaaggcgaca
8101 gcggaaagacc gatcctggac aacagaggca gagttgtggc tattttctaa ggaggtgcaa
8161 atgagggcac gcgtacggcg cttcagttt tcacttggaa ccagaaagggt gtgaccattt
8221 gggataaccccg cgaagggttct gaaccgtggt cactagttt acgcgtatgc gtgcatttgc
8281 atgtcacgtt cccatgcgac aaaccaccccg tgtgttattc actgacgcca gaacgaacac
8341 tcgacgtgtc cgaagagaac gtcgacaatc caaattacga cacgctgctg gagaacgtct
8401 tgaatgtcc atcacgcggg cccaaacgaa gcattaccga tgacttcaca ctgaccagtc
8461 cctacctggg gttctgcggc tattgcagac actcaacgcgt gtgtttcagc ccaataaaaa
8521 ttgagaacgt gtgggacgaa tctgatgatg gatcgattt aatccaggc tcggcacaat
8581 tcggctacaa tcaggcaggc actgcggatg tcaccaaatt ccgttacatg tctttcgacc
8641 acgaccatga catcaaggaa gacagtatgg agaaaaatagc tatcagcaca tctggaccct
8701 gccgtcgct tggccacaaaa gggtaattcc tggtagctca atgtcctcca ggtgacagtg
8761 taaccgtcag tattcacgagc ggagcatctg agaattcatg caccgtggag aaaaagatca
8821 ggaggaagtt tgcggtaga gaggagttact tggccccacc cgtccatgaa aagctggtaa
8881 agtgcacgt ttacgatcac ttgaaggaga cgtctgcggg gtacataacc atgcacaggc
8941 caggcccaca cgcgtataag tcctatctgg aggaagcgtc aggcaagtg tacattaaac
9001 cacttctgg caagaacgtc acctacgaat gtaagtgtgg cgactacagc acaggtatcg
9061 tgagcacgct aacgaagatg aacggctgca ctaaaagcaaa acagtgcatt gcctacaaga
9121 gcgaccaaaac gaaatgggtc ttcaactcgc cggatctttag taggcacaca gaccactcag
9181 tgcaaggtaa attgcacatt ccatccgc tgcacaccgac agtctggcc gttccgttag
9241 ctcacacgct tacagtcaag aagtgggtca aaggcatcac cttccacact actgcaatgc
9301 gaccaacatt gctgacaacg agaaaattgg ggctgcgagc agacgcaaca gcagaatgga
9361 ttacagggtc tacatccagg aattttctg tggggcgaga agggctggag tacgtatggg
9421 gtaaccatga accagtcaag gtctggccccc aggagtgcgc accaggcgac ccacatggat
9481 ggccgcgtga gatcatcatc cactattatc atcggcatcc agtctacact gtcattgtgc
9541 tgcgtgggtc cgctcttgc tgcactggtag gcactgcatt atcagcagct tgcattgcaca

42
DNA
Sequence
Analysis
Program
Version 1.0

9601 aagcaagaag agactgcctg acgccatacg cgcttgcacc gaacgcaacg gtacccacag
9661 cattagcggt tttgtgctgc attcggccaa ccaacgctga aacatttggaa gaaactttga
9721 accatctgtg gtttaacaac caaccgttgc tctggcaca gttgtgcatt cctctggcag
9781 cgcttgttat tctgttccgc tgctttcat gctgcattgc ttttttattt gttgcaggcg
9841 tctgcctggg gaaggttagac gccttcgaac atgcgaccac tgcgcaaat gttccgggg
9901 tcccgatataa ggcgttggc gaacgcgcag gttacgcgc acttaacctg gagatcacgg
9961 tcgtctcattc ggaatttaaca cttcaacta acaaggagta cgtgacctgc aaattccaca
10021 cagtcattcc ttcaccacaa gttaaatgct gcgggtccct cgagtgc当地 gcatcctcaa
10081 aggcggatta cacatgccgc gtttttggcg gtgtgtaccc tttcatgtgg ggaggcgac
10141 aatgcttctg tgacagttag aacacacaac tgagttaggc gtacgtcgag ttcgctccag
10201 actgcactat agatcacgc gtcgcactaa aagttcacac agtgc当地 aaagtccgg
10261 tgcgtatagt atacggcaac accaccgcgc acctggatac gttgtcaat ggcgtcacgc
10321 cagttcctc acgggacctg aaggtcatag cagggccgat atcagccgct tttcaccct
10381 ttgaccataa ggtcgtcatc agaaaaggggc ttgtttacaa ctacgactc cctgagttat
10441 gagctatgaa accaggagcg ttcggcgata ttcaagcatc ctcgcttgc当地 gctacagaca
10501 tagtagcccg cactgacata cggctgctga agccttctgt caagaacatc cacgtccccc
10561 acacccaagc agtatcaggg tatgaaatgt ggaagaacaa ctcaggacga cccctgcaag
10621 aaacagcacc atttggatgt aaaattgaag tggagccctc gcgc当地 aactgtgctt
10681 acgggcacat ccctatctcg attgacatcc ctgatgc当地 ttttgc当地 tc当地
10741 caccacaat tttagaagtt agctgc当地 tagcagactg catttatttgc
10801 gtgggtctct aacattacag tacaagctg acaggagggg acattgtcca gttca
10861 actccacgac agcttttg aaggaagcga ccacacatgt gactgc当地 ggc当地
10921 cactacattt tagcacatcg agcccacaag caaattttgc当地 agttcgct
10981 agtccacctg caatgctgaa tgtaaaccac cggccgacca cataattgga
11041 aagtgcacca agaattccag gccc当地 cccaaacatc ttggactgg ctgc当地
11101 tgtttggggg agcatcatcc ctcattttg taggacttat agtgggctc
11161 tgcttataaa cacacgtaga tgactgagcg cggacactga catagc当地
11221 tacttccgag gaagcgtggc gc当地 acgc当地 tgacactaa
11281 tttccgagga agcacagtgc ataatgctgt gc当地
11341 catattaaca acactatatc actttatgc gactcactat
11401 acatattta cttaaaaaca ctatacacac
11461 ttttatttt gttttaaaaa ttcc
//

SEQUENCE LISTING - pcdwxh-7 - SEQ ID NO 2

<210> SEQ ID NO 2
<211> LENGTH: 4150
<212> TYPE: DNA
<213> ORGANISM: WESTERN EQUINE ENCEPHALITIS VIRUS
<214> FEATURES:

Location/Qualifiers

source 1..4150
/organism="Western equine encephalomyelitis virus"
/strain="71V-1658"
vector sequence 1..9
5' Sac1 primer 9..20
3' end of NS4 gene 16..114
intragenic region 115..159
CDS 160..3869
/note="encodes capsid, E3, E2, 6K and E1 proteins"
/codon_start=160
/product="structural polyprotein"

/translation="MFPYPQLNFPPVYPTNPMAYRDPNPPRRWRPFRPPLAAQIEDL
RRSIVNLTFKQRSPNPPPGPPPKKKSAPKPKPTQPKKKQQAKRTKRKPKGKRQRM
CMKLESDKTFPIMLNGQVNGYACVVGGLMKPLHVEGKIDNEQLAAVKLKKASMYDLE
YGDVPQNMKSDTLQYTSDKPPGFYNWHGAVQYENGRFTVPRGVGGKGDSGRPILDNR
GRVVAIVLGGANEGRTRTALSVVTWNQKGVTIRDTPEGSEPWSLVTALCVLSNVTFC
KPPVCYSLTPERTLDVLEENVDPNYDTLLENVLKCPSSRPKRSITDDFTLTSPYLGF
CPYCRHSTPCFSPIKIENVWDESDDGSIRIQVSAQFGYNQAGTADVTKFYMSFDHDH
DIKEDSMEKIAISTSGPCRRLGHKGYFLLAQCPGDSVTVSITSGASENSCTVEKKIR
RKFVGREELFPPVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
KPPSGKNVTYECKCGDYSTGIVSTRTKMNGCTAKQCIAYKSDQTKWFVFNSPDLIRHT
DHSVQGKLHIPFRLTPTVCPVPLAHTPTVTKWFKGITLHLTAMRPTLLTRKLGLRAD
ATAEWITGSTSRNFSVGREGLEYVWGNHEPVRVWAQESAPGDPHGPHEIIIHYYHRH
PVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT
NAETFGETLNHLWFNNQPFLWAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAF
EHATTVPNVPGI PYKALVERAGYAPLNLEITVVSSELTPSTNKEYVTCKFHTVIPSQ

VKCCGSLECKASSKADYTCRVFGGVYPMWGGACFCDSNTQLSEAYVEFAPDCTID

HAVALKVHTAALKVGLRIVYGNTTAHLDFTVNGVTPGSSRDLKVIAGPISAAFSPFDH

KVVIRKGLVNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVPY

TQAVSGYEMWKNNSGRPLQETAPFGCKIEVEPLRASNCAYGHIPISIDIPDAAFVRSS

ESPTILEVSCTVADCIYSADFGGSLTLQYKADREGHCPVHSHTTAVLKEATTHVTAV

GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHIIGEPHKVDQEFQAAVSKTSW
NWLLALFGGASSLIVVGLIVLVCSSMLINTRR"

3' UTR

3870-4150

BASE COUNT ORIGIN

1	ggccctctag	agctcataact	ggcaggcctg	atcatcacgt	ctctgtccac	gttagccgaa
61	agcguttaaga	acttcaagag	cataagaggg	agcccaatca	ccctctacgg	ctgacctaataa
121	taggtgacgt	agtagacacg	cacctaccca	ccggcagaat	gttccatac	cctcagctga
181	actttccacc	agtttaccct	acaaatccga	tggcttacgg	agatccaaac	cctcttaggc
241	gccgctggag	gccgttccgg	cccccgctgg	ctgctcaaata	cgaagatctt	aggaggtcga
301	tagtcaactt	gactttcaaa	caacgatcac	ctaattccggcc	gccaggtcca	ccgccaaaga
361	agaagaagag	tgctcttaag	ccaaaaccta	ctcagcctaa	aaagaagaag	cagcaagcga
421	agaggacgaa	acgcaagcct	aaaccaggga	aacgacaaacg	tatgtgtatg	aagttggagt
481	cggacaagac	atttccgatc	atgctgaacg	gccaaatgaa	tggatatgcc	tgcgttgcg
541	gaggaaggct	gatgaaacca	ctccacgttg	aaggaaaaat	tgataatgag	caattagcgg
601	ccgtgaaatt	gaagaaggct	agcatgtacg	acttggagta	cggcgcacgtt	ccccagaaca
661	tgaaatcaga	cacgctgcag	tacaccagcg	acaaaccacc	gggcttctac	aactggcacc
721	acggcgcagt	ccagtatgag	aatgggagat	ttaccgtacc	gagaggagtg	ggcgggaaag
781	gcgacagcgg	aagaccgatc	ctggacaaca	gaggcagagt	tgtggctatt	gttcttaggag
841	gtgcaaata	gggcacgcgt	acggcgctt	cagtggtcac	ttgaaaccag	aaaggggtga
901	ccattaggg	taccccccga	ggttctgaac	cgtggtaact	agttacagcg	ctatgcgtgc
961	tttcgaatgt	cacgttccca	tgcgacaaaac	cacccgtgtg	ctattcactg	acgcccagaac
1021	gaacactcga	cgtgctcga	gagaacgtcg	acaatccaaa	ttacgacacg	ctgctggaga
1081	acgtcttga	atgtccatca	cgccggccca	aacgaagcat	taccgatgac	ttcacactga
1141	ccagtccta	cctggggttc	tgcccgtatt	gcagacactc	aacgcgtgt	ttcagccaa
1201	taaaaattga	gaacgtgtgg	gacgaatctg	atgatggatc	gattagaatc	caggtctcg
1261	cacaattcgg	ctacaatcag	gcaggcactg	cggatgtcac	caaattccgt	tacatgtctt
1321	tcgaccacga	ccatgacatc	aaggaagaca	gtatggagaa	aatagctatc	agcacatctg
1381	gaccctgccc	tcgtcttggc	cacaaagggt	acttcctgtt	agctcaatgt	cctccaggtg
1441	acagtgtaac	cgtcagtatc	acgagcggag	catctgagaa	ttcatgcacc	gtggagaaaa
1501	agatcaggag	gaagtttgc	ggttagagagg	agtaacttgg	cccacccgtc	catggaaagc
1561	tggtaaagt	ccacgttac	gatcaacttga	aggagacgtc	tgccgggtac	ataaccatgc
1621	acaggccagg	cccacacgcg	tataagtctt	atctggagga	agcgtcaggc	gaagtgtaca
1681	ttaaaccacc	ttctggcaag	aacgtcacct	acgaatgtaa	gtgtggcgac	tacagcacag
1741	gtatcgtag	cacgcgaacg	aagatgaacg	gctgcactaa	agcaaaaacag	tgcattgcct
1801	acaagagcga	ccaaacgaaa	tgggtcttca	actcgccgga	tcttattagg	cacacagacc
1861	actcagtgc	agttaaattg	cacattccat	tccgcttgcac	accgacagtc	tgccccgttc
1921	cgtagctca	cacgcttaca	gtcacgaagt	ggttcaaagg	catcaccctc	cacctgactg

1981 caatgcgacc aacattgctg acaacgagaa aattggggct gcgagcagac gcaacagcag
2041 aatggattac agggtctaca tccaggaatt tttctgtggg gcgagaaggg ctggagtacg
2101 tatggggtaa ccatgaacca gtcagagtct gggcccagga gtcggcacca ggcgacccac
2161 atgatggcc gcatgagatc atcatccact attatcatcg gcatccagtc tacactgtca
2221 ttgtgctgtg tgggtcgct cttgctatcc tggtaggcac tgcatcatca gcagcttgca
2281 tcgccaaagc aagaagagac tgcctgacgc catacgcgtc tgcaccgaac gcaacggta
2341 ccacagcatt agcggtttg tgctgcattc gccaaaccaa cgctgaaaca tttggagaaa
2401 ctttgaacca tctgtggttt aacaaccaac cgtttctctg ggcacagttg tgcattcctc
2461 tggcagcgct tggattctg ttccgctgct tttcatgtcg catgcctttt ttattgggt
2521 caggcgtctg cctggggaaag gtagacgcct tcgaacatgc gaccactgtg ccaaattgtc
2581 cggggatccc gtataaggcg ttggtcgaac ggcaggtt cgcgcactt aacctggaga
2641 tcacggtcgt ctcatcgaa ttaacacatt caactaacaa ggagtacgtg acctgcaa
2701 tccacacagt cattcctca ccacaagttt aatgctgcgg gtcctcgag tgcaaggcat
2761 cctcaaaggc ggattacaca tgccgcgtt ttggcggtgt gtacccttc atgtggggag
2821 ggcgcacaatg cttctgtgac agtgagaaca cacaactgag tgaggcgta gtcgagttcg
2881 ctccagactg cactatagat cacgcagtcg cactaaaagt tcacacagct gctctgaaag
2941 tcggcctgcg tatagtatac ggcaacacca cgcgcaccc ggatacggtt gtcaatggcg
3001 tcacgccagg ttccctcacgg gacctgaagg tcatagcagg gccgatatac gccgctttt
3061 cacccttga ccataaggc gtcatcagaa aggggcttgc ttacaactac gactccctg
3121 agtatggagc tatgaaaacca ggagcgttcg gcgatattca agcatcctcg cttgatgcta
3181 cagacatagt agcccccaact gacatacggc tgctgaagcc ttctgtcaag aacatccacg
3241 tcccctacac ccaagcagta tcagggtatg aaatgtggaa gaacaactca ggacgacccc
3301 tgcaagaaac agcaccattt ggatgtaaaa ttgaagtggc gcctctgcga gcgtctaact
3361 gtgcttacgg gcacatccct atctcgattt acatccctga tgcaagttt gtgagatcat
3421 cagaatcacc aacaatttttta gaagttatgc gcacagtgc agactgcatt tattctgcag
3481 actttggtgg ttctctaaca ttacagtaca aagctgacag ggagggacat tgcgcgttc
3541 actcccactc cacgacagct gtttgaagg aagcgaccac acatgtgact gccgtaggca
3601 gcataacact acatttttagc acatcgagcc cacaagcaaa ttttatagtt tcgctatgcg
3661 gcaagaagtc cacctgcaat gctgaatgtt aaccaccggc cgaccacata attggagaac
3721 cacataaaatg cgaccaagaa ttccaggcgg cagttccaa aacatcttgg aactggctgc
3781 ttgcactgtt tggggagca tcatccctca ttgtttagg acttatagtt ttggctgc
3841 gctctatgct tataaacaca cgtagatgac tgagcgcggc cactgacata gcggtaaaac
3901 tcgatgtact tccgaggaag cgtgggtgcatt aatgccacgc gccccttgac actaaaactc
3961 gatgtatttc cgaggaagca cagtgcataa tgctgtgcag tgtcacatta atcgatatac
4021 acactacata ttaacaacac tatacactt ttatgagact cactatgggt ctctaata
4081 cactacacat atttactta aaaacactat acacacttta taaattctct cataattca
4141 cttaggttt
//

SEQUENCE LISTING - pVHX-6 - SEQ ID NO 3

<200> SEQUENCE CHARACTERISTICS
<210> SEQ ID NO 3
<211> LENGTH: 1260
<212> TYPE: DNA
<213> ORGANISM: EXAMPLE OF A CMV PROMOTER COUPLED TO THE pcDW-XH7
SEQUENCE
<220> FEATURE:
1-302 pVAX vector sequence
302-401 CMV promoter
411 CMV putative transcriptional start site
444-463 T7 promoter
464-509 pVAX multicloning region
510 start of pcDW-HX45 sequence

1 cattgacgtc aatgggtgga gtatttacgg taaaactgcc acttggcagt acatcaagtg
61 tattcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgccctggcat
121 tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta cgtatttagtc
181 atcgcttta ccatgggtgat gcggttttgg cagtagatca atggcgtgg atagcggtt
241 gactcacggg gatttccaag tctccacccc attgacgtca atggaggtt gtttggcac
301 caaaatcaac gggactttcc aaaatgtcgt aacaactccg ccccattgac gcaaatggc
361 ggtaggcgtg tacgggtgggaa ggtctatata agcagagctc tctggctaac tagagaaccc
421 actgcttact ggcttatacg aattaatacg actcactata gggagaccca agctggctag
481 cgtttaaact taagcttggt accgagctca tactggcagg cctgatcatc acgtctctgt
541 ccacgttagc cgaaagcggtt aagaacttca agagcataag agggagccca atcaccctct
601 acggctgacc taaataggtg acgttagtata cacgcaccta cccaccgcca gaatgtttcc
661 ataccctcag ctgaactttc caccagtttta ccctacaaat ccgatggctt accgagatcc
721 aaaccctcct aggcgcccgtt ggaggccgtt tcggcccccg ctggctgctc aaatcgaaga
781 tcttaggagg tcgatagtca acttgacttt caaacaacga tcacctaatt cggccgcagg
841 tccacccgcca aagaagaaga agagtgcctc taagccaaa cctactcagc ctaaaaagaa
901 gaagcagcaa gccaaagagga cgaaacgc当地 gcctaaacca gggaaacgc当地 aacgtatgtg
961 tatgaagttt gatcgccgaca agacatttcc gatcatgctg aacggccaa tgaatggata
1021 tgcctgcgtt gtcggagggaa ggctgatgaa accactccac gttgaaggaa aaattgataa
1081 tgagcaatta gcgccgtga aattgaagaa ggcttagcatg tacgacttgg agtacggcga
1141 cgttccccag aacatgaaat cagacacgct gcagtagtacc agcgacaaac caccgggctt
1201 ctacaactgg caccacggcg cagttccagta tgagaatggg agatttaccg taccgagagg

. CONTINUOUS WITH pcDW-HX45 (pcDW-XH7) SEQUENCE

4177 gaatgtaaac caccggccga ccacataatt ggagaaccac ataaagtca ccaagaattc
4237 caggccggcag tttccaaaac atcttggAAC tggctgcttgc cactgtttgg gggagccatca
4297 tccctcatttgc ttgttaggact tatagtgttgc gtctgcagct ctatgctttaa aacacacgt
4357 agatgactga gcgccggacac tgacatagcg gtaaaaactcg atgtacttcc gaggaagcgt
4417 ggtgcataat gccacgc当地 gcttgacact aaaactcgat gtatttccga ggaaggccac
4477 tgcataatgc tgcgtcactgtt cacattaatc gtatatacaca ctacatatttaca acaacactat
4537 atcactttta tgagactcac tatgggtctc taatatacac tacacatattttaa ttactttaaaa
4597 acactataca cactttataa attcttttat aatttttctt ttgctttaga gcacactggc

4657 ggccgttact agtggatccg agctctagag ggcc

4177-4643 pcDW-HX45
4644-4790 vector sequence //